

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Bandman, Olga
Goli, Surya K.
Hillman, Jennifer L.

(ii) TITLE OF THE INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE

(iii) NUMBER OF SEQUENCES: 12

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
(B) STREET: 3174 Porter Drive
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: To Be Assigned
(B) FILING DATE: Herewith
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Billings, Lucy J.
(B) REGISTRATION NUMBER: 36,749
(C) REFERENCE/DOCKET NUMBER: PF-0187 US

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 415-855-0555
(B) TELEFAX: 415-845-4166

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 264 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: Consensus
(B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met Ala Ala Ala Ala Val Ala Arg Leu Trp Trp Arg Gly Ile Leu Gly
 1 5 10 15
 Ala Ser Ala Leu Thr Arg Gly Thr Gly Arg Pro Ser Val Leu Leu Leu
 20 25 30
 Pro Val Arg Arg Glu Ser Ala Gly Ala Asp Thr Arg Pro Thr Val Arg
 35 40 45
 Pro Arg Asn Asp Val Ala His Lys Gln Leu Ser Ala Phe Gly Glu Tyr
 50 55 60
 Val Ala Glu Ile Leu Pro Lys Tyr Val Gln Gln Val Gln Val Ser Cys
 65 70 75 80
 Phe Asn Glu Leu Glu Val Cys Ile His Pro Asp Gly Val Ile Pro Val
 85 90 95
 Leu Thr Phe Leu Arg Asp His Thr Asn Ala Gln Phe Lys Ser Leu Val
 100 105 110
 Asp Leu Thr Ala Val Asp Val Pro Thr Arg Gln Asn Arg Phe Glu Ile
 115 120 125
 Val Tyr Asn Leu Leu Ser Leu Arg Phe Asn Ser Arg Ile Arg Val Lys
 130 135 140
 Thr Tyr Thr Asp Glu Leu Thr Pro Ile Glu Ser Ala Val Ser Val Phe
 145 150 155 160
 Lys Ala Ala Asn Trp Tyr Glu Arg Glu Ile Trp Asp Met Phe Gly Val
 165 170 175
 Phe Phe Ala Asn His Pro Asp Leu Arg Arg Ile Leu Thr Asp Tyr Gly
 180 185 190
 Phe Glu Gly His Pro Phe Arg Lys Asp Phe Pro Leu Ser Gly Tyr Val
 195 200 205
 Glu Leu Arg Tyr Asp Asp Glu Val Lys Arg Val Val Ala Glu Pro Val
 210 215 220
 Glu Leu Ala Gln Glu Phe Arg Lys Phe Asp Leu Asn Ser Pro Trp Glu
 225 230 235 240
 Ala Phe Pro Val Tyr Arg Gln Pro Pro Glu Ser Leu Lys Leu Glu Ala
 245 250 255
 Gly Asp Lys Lys Pro Asp Ala Lys
 260

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1023 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Consensus
- (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GAACTCTAAT	ACGAGCACTA	TAGGGAAAGC	TGGTAGCCTG	CAGGTACCGG	TCCGGAATTC	60
CCGGGTCGAC	CCACGCGTCC	GCCGTGCCCT	TGGGGCTCCG	TGTCCTGCTG	TCTTTCCGTC	120
CGCTGCCTAG	TCTGCATCTG	AGTAACATGG	CGGCGGCGGC	GGTAGCCAGG	CTGTGGTGGC	180
GCGGGATCTT	GGGGGCCTCG	GCGCTGACCA	GGGGGACTGG	GCGACCCTCC	GTTCTGTTGC	240
TGCCGGTGAG	GCGGGAGAGC	GCCGGGGCCG	ACACGCGCCC	CACTGTCAGA	CCACGGAATG	300
ATGTGGCCCA	CAAGCAGCTC	TCAGCTTTTG	GAGAGTATGT	GGCTGAAATC	TTGCCCAAGT	360
ATGTCCAACA	AGTTCAGGTG	TCCTGCTTCA	ATGAGTTAGA	GGTCTGTATC	CATCCTGATG	420
GCGTCATCCC	AGTGCTGACT	TTCCCTCAGGG	ATCACACCAA	TGCACAGTTC	AAATCTCTGG	480
TTGACTTGAC	AGCAGTGGAC	GTCCCAACTC	GGCAAAACCG	TTTTGAGATT	GTCTACAACC	540
TGTTGTCTCT	GCGCTTCAAC	TCACGGATCC	GTGTGAAGAC	CTACACAGAT	GAGCTGACGC	600
CCATTGAGTC	TGCTGTCTCT	GTGTTCAAGG	CAGCCAAC TG	GTATGAAAGG	GAGATCTGGG	660
ACATGTTTGG	AGTCTTCTTT	GCTAACCACC	CTGATCTAAG	AAGGATCCTG	ACAGATTATG	720

GCTTCGAGGG	ACATCCTTTC	CGGAAAGACT	TTCCTCTATC	TGGCTATGTT	GAGTTACGTT	780
ATGATGATGA	AGTGAAGCGT	GTGGTGGCAG	AGCCGGTGGA	GTTGGCCCAA	GAGTTCCGCA	840
AATTTGACCT	GAACAGCCCC	TGGGAGGCTT	TCCCAGTCTA	TCGCCAACCC	CCGGAGAGTC	900
TCAAGCTTGA	AGCCGGAGAC	AAGAAGCCTG	ATGCCAAGTA	GCTCCAGGGA	ACGCATGTGG	960
ATCCTAGACA	GCGCCTTATC	TATGATTGAG	TGTCCGTGTA	AATAAATTCC	TACTTAGACT	1020
TAC						1023

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 129 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Consensus
- (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Ser	Phe	Pro	Lys	Tyr	Lys	Pro	Ser	Ser	Leu	Arg	Thr	Leu	Pro	Glu
1				5					10					15	
Thr	Leu	Asp	Pro	Ala	Glu	Tyr	Asn	Ile	Ser	Pro	Glu	Thr	Arg	Arg	Ala
			20				25					30			
Gln	Ala	Glu	Arg	Leu	Ala	Ile	Arg	Ala	Gln	Leu	Lys	Arg	Glu	Tyr	Leu
		35				40					45				
Leu	Gln	Tyr	Asn	Asp	Pro	Asn	Arg	Arg	Gly	Leu	Ile	Glu	Asn	Pro	Ala
	50					55				60					
Leu	Leu	Arg	Trp	Ala	Tyr	Ala	Arg	Thr	Ile	Asn	Val	Tyr	Pro	Asn	Phe
65				70				75						80	
Arg	Pro	Thr	Pro	Lys	Asn	Ser	Leu	Met	Gly	Ala	Leu	Cys	Gly	Phe	Gly
			85					90					95		
Pro	Leu	Ile	Phe	Ile	Tyr	Tyr	Ile	Ile	Lys	Thr	Glu	Arg	Asp	Arg	Lys
		100					105						110		
Glu	Lys	Leu	Ile	Gln	Glu	Gly	Lys	Leu	Asp	Arg	Thr	Phe	His	Leu	Ser
		115					120					125			
Tyr															

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 451 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Consensus
- (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCAAGATGTC	GTTCCCAAAG	TATAAGCCGT	CGAGCCTGCG	CACTCTGCCT	GAGACCCTCG	60
ACCCAGCCGA	ATACAACATA	TCTCCGGAAG	CCCGGCGGGC	GCAAGCGGAG	CGGTTGGCCA	120
TAAGAGCCCA	GCTGAAACGA	GAGTACCCTG	TTCAGTACAA	CGATCCCAAC	CGCCGAGGGC	180
TCATCGAAAA	TCCTGCCTTG	CTTCGTTGGG	CCTATGCAAG	AACAATAAAT	GTCTATCCTA	240
ATTTTCAGACC	CACTCCTAAA	AACTCACTCA	TGGGAGCTCT	GTGTGGATTT	GGGCCCCCTCA	300
TCTTCATTTA	TTATATTATC	AAAACCTGAG	GGGATAGGAA	AGAAAAACTT	ATCCAGGAAG	360
GAAAATTGGA	TCGAACATTT	CACCTCTCAT	ATTAAGTCTG	GCAATGATGA	CTATATGTAT	420

TCCTGCCTAA ATAAATCATC TATTAATCAT T

451

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 106 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Consensus
- (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Pro	Phe	Leu	Asp	Ile	Gln	Lys	Arg	Phe	Gly	Leu	Asn	Ile	Asp	Arg
1				5					10					15	
Trp	Leu	Thr	Ile	Gln	Ser	Gly	Glu	Gln	Pro	Tyr	Lys	Met	Ala	Gly	Arg
			20					25					30		
Cys	His	Ala	Phe	Glu	Lys	Glu	Trp	Ile	Glu	Cys	Ala	His	Gly	Ile	Gly
		35					40					45			
Tyr	Thr	Arg	Ala	Glu	Lys	Glu	Cys	Lys	Ile	Glu	Tyr	Asp	Asp	Phe	Val
	50					55				60					
Glu	Cys	Leu	Leu	Arg	Gln	Lys	Thr	Met	Arg	Arg	Ala	Gly	Thr	Ile	Arg
65					70					75				80	
Lys	Gln	Arg	Asp	Lys	Leu	Ile	Lys	Glu	Gly	Lys	Tyr	Thr	Pro	Pro	Pro
			85						90					95	
His	His	Ile	Gly	Lys	Gly	Glu	Pro	Arg	Pro						
			100					105							

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 470 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Consensus
- (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGCTAGTCGT	TCTGAAGCGG	CGGCCAGAGA	AGAGTCAAGG	GCACGAGCAT	CGGCCATGCC	60
TTTCTTGAC	ATCCAGAAAA	GGTTCGGCCT	TAACATAGAT	CGATGGTTGA	CAATCCAGAG	120
TGGTGAACAG	CCCTACAAGA	TGGCTGGTCG	ATGCCATGCT	TTTGAAAAAG	AATGGATAGA	180
ATGTGCACAT	GGAATCGGTT	ATACTCGGGC	AGAGAAAGAG	TGCAAGATAG	AATATGATGA	240
TTTCGTAGAG	TGTTTGCTTC	GGCAGAAAAC	GATGAGACGT	GCAGGTACCA	TCAGGAAGCA	300
GCGGGATAAG	CTGATAAAGG	AAGGAAAGTA	CACCCCTCCA	CCTCACCACA	TTGGCAAGGG	360
GGAGCCTCGG	CCCTGAACAG	AGCAGCTGCT	GATGTCTGGA	GGCTGATTTT	CCTGTTCTCT	420
GTTCTCCACT	GGAAAGGTTG	TTTACGACAA	ACCTCCTTGT	CAAAGTGTGT		470

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (vii) IMMEDIATE SOURCE:
 (A) LIBRARY: Consensus
 (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met	Ile	Ala	Arg	Asn	Pro	Glu	Pro	Leu	Arg	Phe	Leu	Pro	Asp	Glu
1			5					10					15	
Ala	Arg	Ser	Leu	Pro	Pro	Lys	Leu	Thr	Asp	Pro	Arg	Leu	Leu	Tyr
			20				25					30		
Ile	Gly	Phe	Leu	Gly	Tyr	Cys	Ser	Gly	Leu	Ile	Asp	Asn	Leu	Arg
		35				40					45			
Arg	Arg	Pro	Ile	Ala	Thr	Ala	Gly	Leu	His	Arg	Gln	Xaa	Xaa	Tyr
		50				55					60			
Thr	Ala	Phe	Phe	Phe	Ala	Gly	Tyr	Tyr	Xaa	Val	Lys	Arg	Glu	Asp
65					70				75					80
Leu	Tyr	Ala	Val	Arg	Asp	Arg	Glu	Met	Phe	Gly	Tyr	Met	Lys	Leu
				85					90					95
Pro	Glu	Asp	Phe	Pro	Glu	Glu	Asp	Lys	Lys	Thr	Tyr	Gly	Glu	Ile
			100					105					110	
Glu	Lys	Phe	His	Pro	Ile	Arg								
														115

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 576 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (vii) IMMEDIATE SOURCE:
 (A) LIBRARY: Consensus
 (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGCAGAGGAG	GAGGAGAAAAG	CTGACCGCTT	AGGCCCGGGT	AGTGGTCGTC	GTGGTTTTTC	60
TTGTAGTTCG	TGGTCTGAGA	CCAGGCCTCA	AGTGGAACG	GCGTCACCAT	GATCGCACGG	120
CGGAACCCAG	AACCCTTACG	GTTTCTGCCG	GATGAGGCCC	GGAGCCTGCC	CCCGCCCAAG	180
CTGACCGACC	CGCGGCTCCT	CTACATCGGC	TTCTTGGGCT	ACTGCTCCGG	CCTGATTGAT	240
AACCTGATCC	GGCGGAGGCC	GATCGCGACG	GCTGGTTTGC	ATCGCCAGNT	TNTATATATT	300
ACGGCCTTTT	TTTTTGCTGG	ATATTATNTT	GTAAAACGTG	AAGACTACCT	GTATGCTGTG	360
AGGGACCGTG	AAATGTTTGG	ATATATGAAA	TTACATCCAG	AGGATTTTCC	TGAAGAAGAT	420
AAGAAAACAT	ATGGTGAAAT	TTTTGAAAAA	TTCCATCCAA	TACGTTGAAG	TCTTCAAAAT	480
GCTTGCTCCA	GTTTCACTGA	TACCTGCTGT	TTCTGAATTT	GATGGAACAT	GTTTCTTATG	540
ACAGTTGAAG	CTTATGCTAA	TCTGTATGTT	GACACC			576

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 266 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (vii) IMMEDIATE SOURCE:
 (A) LIBRARY: GenBank
 (B) CLONE: 163416

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```

Met Ala Ala Ala Val Ala Ala Ala Ala Pro Gly Cys Trp Gln Arg Leu
 1      5      10
Val Gly Ser Ala Pro Ala Arg Val Ala Gly Arg Pro Ser Val Leu
      20      25      30
Leu Leu Pro Val Arg Arg Glu Ser Ser Ala Ala Asp Thr Arg Pro Thr
      35      40      45
Val Arg Pro Arg Asn Asp Val Ala His Lys Gln Leu Ser Ala Phe Gly
      50      55      60
Glu Tyr Val Ala Glu Ile Leu Pro Lys Tyr Val Gln Gln Val Gln Val
      65      70      75      80
Ser Cys Phe Asn Glu Leu Glu Ile Cys Ile His Pro Asp Gly Val Ile
      85      90      95
Pro Val Leu Thr Phe Leu Arg Asp His Ser Asn Ala Gln Phe Lys Ser
      100      105      110
Leu Ala Asp Leu Thr Ala Val Asp Ile Pro Thr Arg Gln Asn Arg Phe
      115      120      125
Glu Ile Val Tyr Asn Leu Leu Ser Leu Arg Phe Asn Ser Arg Ile Arg
      130      135      140
Val Lys Thr Tyr Thr Asp Glu Leu Thr Pro Ile Glu Ser Ser Val Pro
      145      150      155      160
Val Tyr Lys Ala Ala Asn Trp Tyr Glu Arg Glu Ile Trp Asp Met Phe
      165      170      175
Gly Val Phe Phe Ala Asn His Pro Asp Leu Arg Arg Ile Leu Thr Asp
      180      185      190
Tyr Gly Phe Glu Gly His Pro Phe Arg Lys Asp Phe Pro Leu Ser Gly
      195      200      205
Tyr Val Glu Leu Arg Tyr Asp Asp Glu Val Lys Arg Val Val Ala Glu
      210      215      220
Pro Val Glu Leu Ala Gln Glu Phe Arg Lys Phe Asp Leu Asn Ser Pro
      225      230      235      240
Trp Glu Ala Phe Pro Ala Tyr Arg Gln Pro Pro Glu Ser Leu Lys Leu
      245      250      255
Glu Ala Gly Asp Thr Lys Pro Glu Ala Lys
      260      265

```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 129 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 114

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

Met Ser Phe Pro Lys Tyr Glu Ala Ser Arg Leu Ser Ser Leu Pro Thr
 1      5      10      15
Thr Leu Asp Pro Ala Glu Tyr Asp Ile Ser Ser Glu Thr Arg Lys Ala
      20      25      30
Gln Ala Glu Arg Leu Ala Ile Arg Ser Arg Leu Lys Arg Glu Tyr Gln
      35      40      45
Leu Gln Tyr Tyr Asp Pro Ser Arg Arg Gly Val Ile Glu Asp Pro Ala
      50      55      60
Leu Val Arg Trp Thr Tyr Ala Arg Ser Ala Asn Ile Tyr Pro Asn Phe
      65      70      75      80
Arg Pro Asn Thr Lys Thr Ser Leu Leu Gly Ala Leu Phe Gly Ile Gly
      85      90      95

```

Pro Leu Val Phe Trp Tyr Tyr Val Phe Lys Thr Asp Arg Asp Arg Lys
 100 105 110
 Glu Lys Leu Ile Gln Glu Gly Lys Leu Asp Arg Thr Phe Asn Ile Ser
 115 120 125
 Tyr

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 106 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 224

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Pro Phe Phe Asp Val Gln Lys Arg Leu Gly Val Asp Leu Asp Arg
 1 5 10 15
 Trp Met Thr Ile Gln Ser Ala Glu Gln Pro His Lys Ile Pro Ser Arg
 20 25 30
 Cys His Ala Phe Glu Lys Glu Trp Ile Glu Cys Ala His Gly Ile Gly
 35 40 45
 Ser Ile Arg Ala Glu Lys Glu Cys Lys Ile Glu Phe Glu Asp Phe Arg
 50 55 60
 Glu Cys Leu Leu Arg Gln Lys Thr Met Lys Arg Leu His Ala Ile Arg
 65 70 75 80
 Arg Gln Arg Glu Lys Leu Ile Lys Glu Gly Lys Tyr Thr Pro Pro Pro
 85 90 95
 His His Ser Gly Gln Glu Glu Pro Arg Ser
 100 105

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 582

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Met Thr Gly Arg Gln Gly Arg Ala Thr Phe Gln Phe Leu Pro Asp
 1 5 10 15
 Glu Ala Arg Ser Leu Pro Pro Pro Lys Leu Thr Asp Pro Arg Leu Ala
 20 25 30
 Phe Val Gly Phe Leu Gly Tyr Cys Ser Gly Leu Ile Asp Asn Ala Ile
 35 40 45
 Arg Arg Arg Pro Val Leu Leu Ala Gly Leu His Arg Gln Leu Leu Tyr
 50 55 60
 Ile Thr Ser Phe Val Phe Val Gly Tyr Tyr Leu Leu Lys Arg Gln Asp
 65 70 75 80
 Tyr Met Tyr Ala Val Arg Asp His Asp Met Phe Ser Tyr Ile Lys Ser

